

SuperStore US 2015 Data Analysis.
A Capstone Project Report in Data Analytics

By
TERUNGWA, Emmanuel

In partial fulfillment for the award of Data Analytics Certification at Techcrush

JUN, 2025

TABLE OF CONTENT

Cover page	1
Table of Content	2
Abstract	3
Acknowledgement	4
Introduction	5
Problem Statement	5
Dataset Overview	6
Project Objectives	6
Project Aim	6
Project Task	7
Business Insights and Recommendations	13
Project Limitations	15
Summary	15
Conclusions	16

ABSTRACT

This project presents an analytical exploration of the SuperStore US 2015 dataset to uncover key insights into sales performance, customer behavior, and regional trends. The dataset comprises transactional records including product categories, sales figures, profit margins, order dates, customer segments, and geographical locations. Using data visualization and analytical advance Excel, we analyzed the distribution of sales and profit by region, category, and sub-category. Our findings indicate that while some regions, like East, West generate high sales and profit, others like central and South show high sales but low profitability, suggesting potential inefficiencies or pricing issues. Segment analysis reveals that corporate customers contribute significantly to revenue, while the home office segment exhibits lower profitability. The technology category emerged as the most profitable, while furniture showed lower returns. These insights can support strategic decisions on inventory management, regional marketing, and customer targeting. Overall, the analysis underscores the value of data-driven decision-making in optimizing business performance and improving profitability across different dimensions of the retail landscape.

Acknowledgement

I would like to express my sincere gratitude to God Almighty for His grace and strength that enabled me to successfully complete this programme.

Special thanks go to the Student Success Executive (SSE), Ms. Mary Oladokun, as well as my tutors and erudite data analytics wizards, Mrs. Grace Emeruwa and Mr. Olaniyi Eluwole, whose tireless support and dedication contributed immensely to this achievement. My appreciation also goes to the entire staff of Techcrush and my students colleagues for their prayers and encouragement.

Finally, I extend my heartfelt appreciation to the CEO and the entire team at TeamO Digital Solutions for their unwavering support in ensuring I completed this programme.

1.0 Introduction:

This report presents a comprehensive data analysis of the SuperStore US 2015 sales data. The primary objectives are to uncover key insights into sales performance, customer behavior, regional trends, and the impact of various factors like discounts and shipping cost on profitability. By leveraging data visualization and analytical techniques, this project aims to provide actionable recommendations to optimize business strategies, improve profitability, and enhance overall operational efficiency.

The SuperStore US 2015 dataset is a rich source of transactional records, encompassing product categories, sales figures, profit margins, orders dates, customer segments, and geographical locations. The analysis covers different aspects of the business, including product performance, regional contributions, and customer purchasing behavior.

The ultimate aim is to provide the SuperStore with a data-driven foundation for future growth and sustained success.

1.1 Problem Statement:

The SuperStore, like many retail businesses, faces challenges in optimizing its sales and profit performance. Without a clear understanding of which products, categories, regions, or customer segments are performing well and which are underperforming, it is difficult to make informed business decisions. This lack of clear insight can lead to:

- i. Inefficient resource allocation: Marketing efforts and inventory investments might be misdirected towards less profitable areas.
- ii. Missed opportunities: High-potential products or regions might not receive the attention they deserve, leading to lost revenue.
- iii. Unaddressed losses: Areas with negative profit margins might continue to drain resources without being identified and corrected.
- iv. Suboptimal strategies: Business strategies might be based on assumptions rather than data-driven facts, leading to ineffective outcomes.

Specifically, the SuperStore needs to address questions such as: Which product categories are the most profitable? Are there any regions consistently losing money? Which customer segments are most valuable? How do discounts affect profit margins? Answering these questions requires a systematic analysis of the available sales data. The core problem is the need for actionable insights derived from raw data to drive strategic improvements and ensure sustainable growth.

1.2 Dataset Overview:

The dataset consists of transactional sales records from a superstore. It contains customer details, order information, shipping details, and sales data. Below are the key attributes of the dataset:

Order Details: Order ID, Order Date, Shipping Cost, Discount

Customer Information: Customer ID, Name, Segment (Corporate, Small Business, etc)

Product Information: Product Category, Sub-Category, Container Type, Unit Price

Financial Metrics: Quantity Sold, Sales, Profit, Product Margin.

Geographical Data: Country, Region, State, City, Postal Code.

1.3 Objectives

Below are the following objectives:

1. Data Clean and preprocessing; Ensuring data integrity and preparing the dataset for analysis.
2. Analyze overall sales and profit performance; Understanding the general trends in sales and profit.
3. Identify best-selling and worst-performing products; Pinpointing products that drives success or face challenges.
4. Investigate the impact of discount on profitability; Analyzing how discounts affect profit margins.
5. Examine the relationship between shipping costs and profit; Understanding the influence of shipping expenses on profitability.
6. Evaluate regional and state-wise sales trends; Assessing geographical performance.
7. Assess which customer segments contribute the most to revenue; Identifying valuable customer groups.
8. Identify inefficiencies in the supply chain or product performance; Highlighting areas for operational improvement.

1.4 Aim

The primary aim of this project is to provide the SuperStore_US 2025 with actionable, data-driven insights to optimize its sales strategies, improve profitability, and enhance overall business efficiency based on the 2015 sales data.

1.5 Project Task:

Task 1: Data Cleaning and Preprocessing:

First, the Excel file was loaded into a Excel Power query. An initial inspection was performed to understand the dataset's structure, including checking data types, identifying the number of rows and columns, handling missing values in product base margin, standardized inconsistent values in Order Priority, Calculated order processing time between order date and shifting date, created calculated fields for profit margins and discount impact. This step confirmed that the CLEANED DATA sheet in excel file contained 1952 entries and 39 columns, with a mix of numerical and

object (text) data types. Crucially, it was confirmed that there were no missing values in the key columns used for analysis, ensuring data integrity for subsequent steps which addressed objective one above.

Task 2: Exploratory Data Analysis (EDA)

Sales and Profit Analysis. The analysis of sales and profit across different product categories reveals significant variations. The three main product categories are Technology, Furniture, and Office Supplies. Each category contributes differently to the overall sales and profit of the SuperStore. The chart below explains sales and profit analysis by product category.

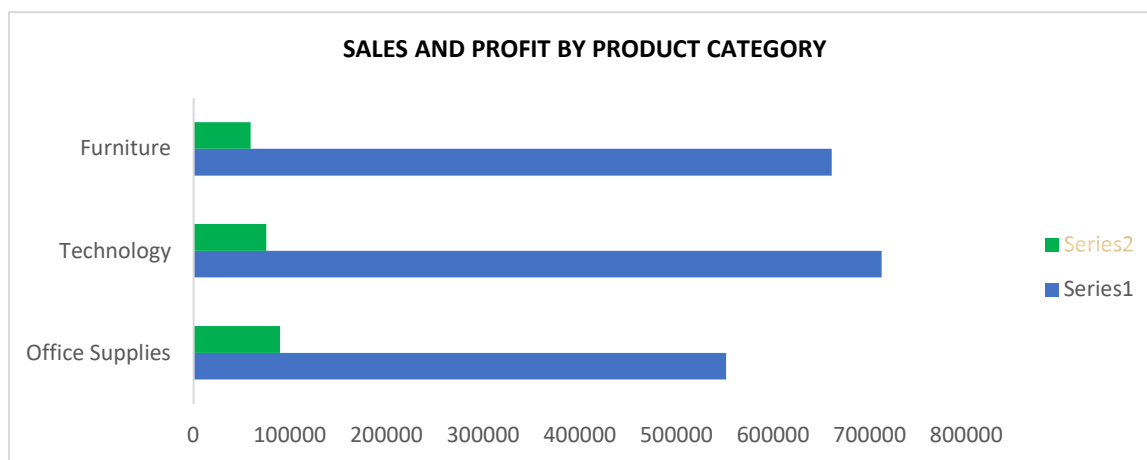


Fig 1. *Sales and profit by product category*

As observed from the fig 1 above, the blue and green colors in the chart represent sales and profit margins respectively. Technology leads in total sales, followed by Furniture, and then Office Supplies. This indicates that customers are spending the most on technology related products.

When examining profit, Office supplies surprisingly emerge as the most profitable category, despite having lower sales than Technology and Furniture. This suggests that Office Supplies have higher profit margins or lower associated costs. Technology is the second most profitable, while Furniture, despite its high sales, yields the lowest profit among the three, indicating potential issues with its profit margins or higher operational costs.

1. Sales and Profit by Product Sub-Category.

Delving deeper into product performance, we analyzed sales and profit at the sub-category level. This analysis view helps pinpoint specific products that are driving success or facing challenges.

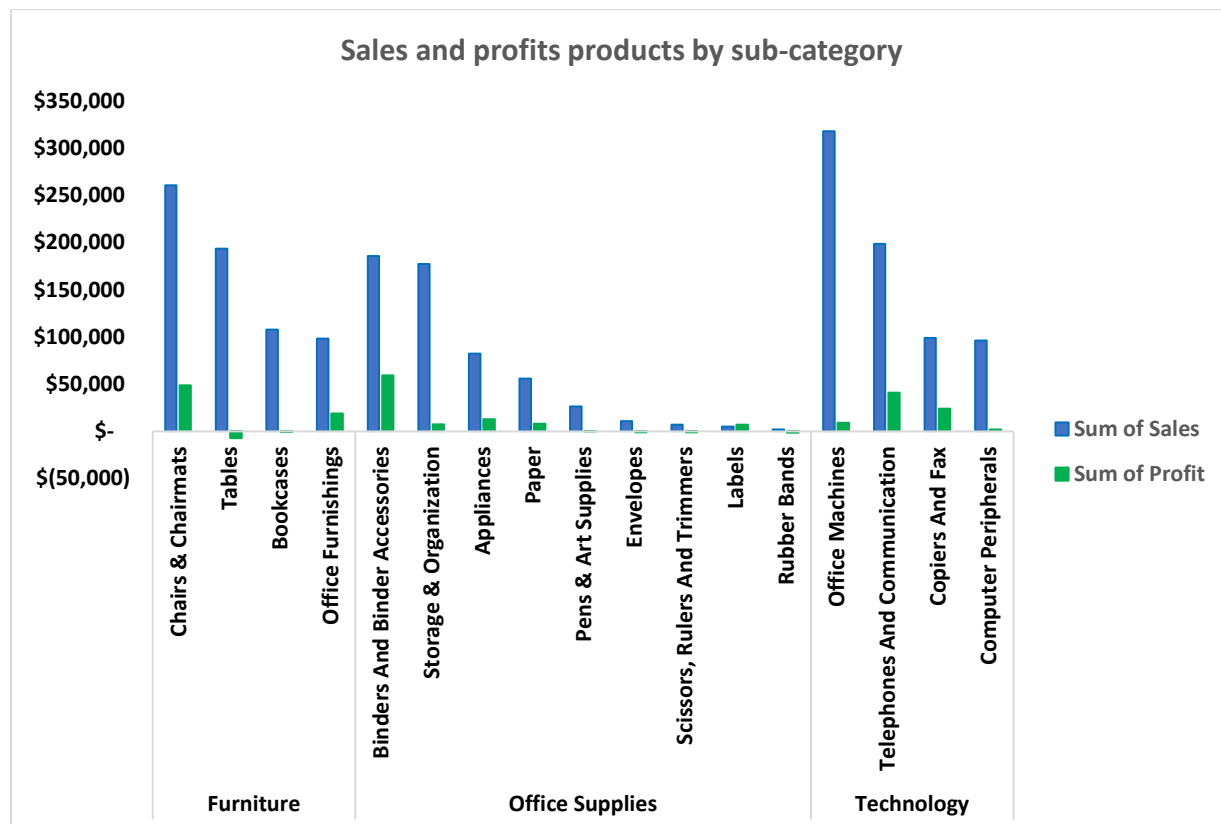


Fig 2. Sales and Profits Product by sub-category.

The sub-categories by sales and profit are dominated by office machines, chairs and chairmats, and telephone and communication. These sub-categories are significant revenue generators for the SuperStore. While sub-categories by sales like rubber bands, and Scissors, Rulers and Trimmers contribute minimally to overall sales and might require strategies to boost their performance. In terms of profit, Binders and Binder Accessories, chairs and chairmats, and Telephones and communication are among the most profitable sub-categories. This indicates healthy profit margins and efficient operations within these areas. However, several sub-categories are experiencing significant losses, including Rubber bands, Envelopes, and Pens & Art Supplies. These negative profit (Loss) figures highlight areas where the superstore is losing, potentially due to high costs, low selling prices, or excessive discounts. Immediate attention is needed to address these underperforming sub-categories.

2. Sales and Profit by Region.

Geographical analysis is crucial for understanding regional performance and identifying areas for targeted strategies. The SuperStore operates across four main regions: East, West, Central, and South. The East and West regions lead in sales, contributing the most to the SuperStore's revenue. The Central region follows, while the South region has the lowest sales volume.

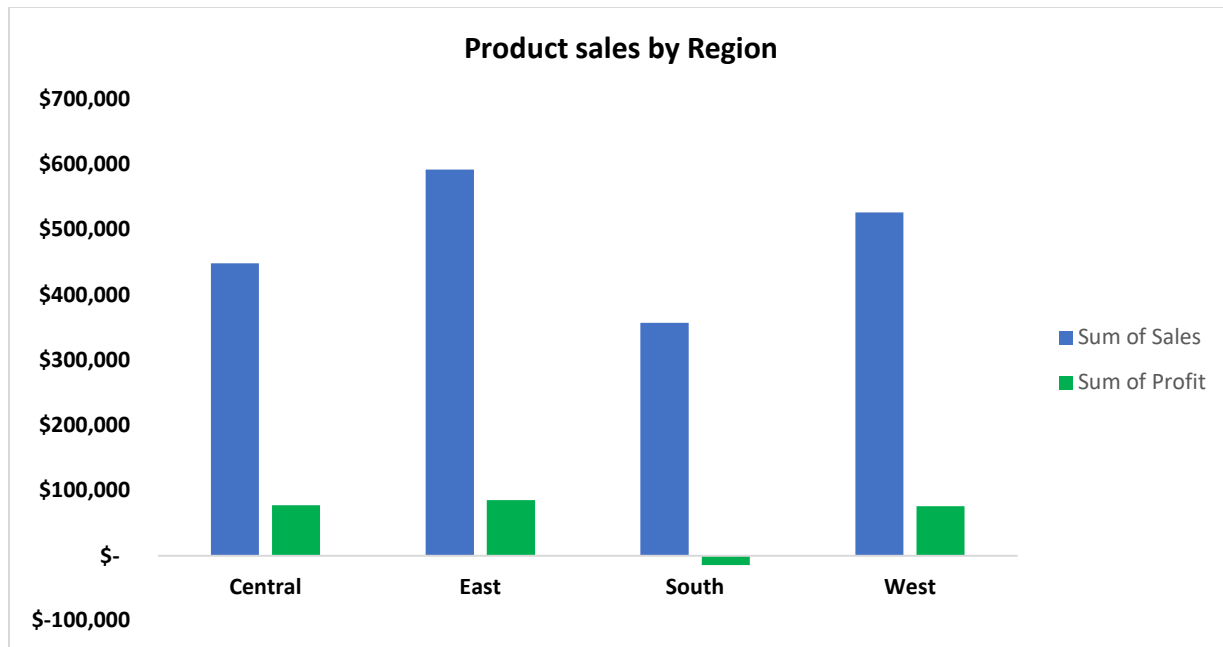


Fig 3. Sales and Profit by Region

In terms of profit, the East, West, and Central regions all show positive profitability, with the East region being the most profitable. However, the South region exhibits a significant loss, indicating a critical area for intervention. This negative profit (loss) in the South region could be due to various factors such as high operational costs, intense competition, or unfavorable pricing strategies.

3. Profit by Customer Segment:

All segments contributed positively to profit. Notably, the Small Business segment, despite lower sales, showed a strong profit contribution, suggesting high profitability per transaction within this segment. Corporate and Home Office also contributed significantly, while Consumer had the lowest profit contribution among the profitable segments.

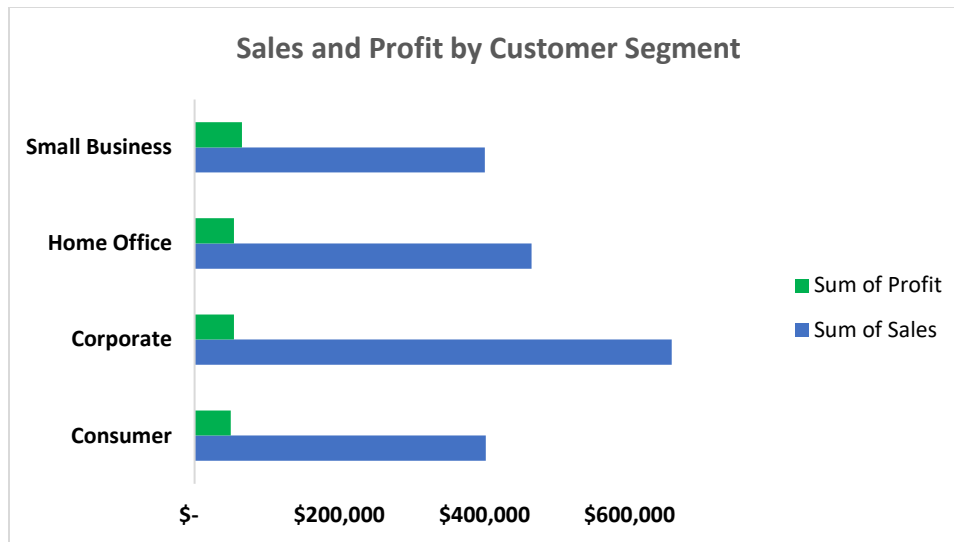


Fig 4. Sales and Profit by Customer segment

This comprehensive analysis provides a clear picture of the superstore’s performance across various dimensions, laying the groundwork for actionable insights and recommendations.

4. Impact of Discounts on sales and Profit.

This analysis investigated how different discount levels affect sales and profitability, crucial for optimizing pricing strategies.

	Discount	of Profit_Margin(%)
Discount	1	
Sum of Profit_Margin	0.147904543	1

Table 1

The analysis revealed that a weak positive correlation (0.147904543) between discount rate and profit margins. This suggest that simply increasing discounts does not proportionally increase profit.

The critical finding is the identification of a discount threshold (0.17 or 17%) beyond which discounts tend to become unprofitable. This implies that discounts exceeding this level can erode profit margins, even if they drive higher sales volume. The insights emphasize the need for strategic discounting, where promotions are carefully evaluated to ensure they contribute positively to the bottom line rather than just increasing sales volume at the expense of profit.

The table below shows the threshold at which discount becomes unprofitable in red text format.

Discount	Sum of Profit
0.03	\$ 59,886
0.02	\$ 42,717
0.05	\$ 36,109
0.06	\$ 31,026
0	\$ 29,472
0.01	\$ 23,015
0.07	\$ 12,959
0.08	\$ 7,326
0.1	\$ 4,284
0.17	\$ (9)
0.21	\$ (18)
0.04	\$ (4,666)
0.09	\$ (18,025)
Grand Total	\$ 224,078

Table 2

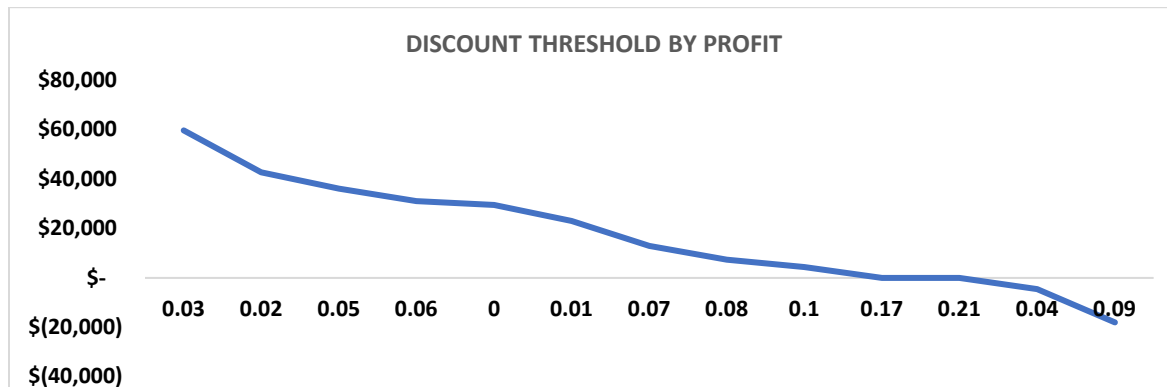


Fig 5 Line chart showing threshold at which discount become unprofitable at 0.17 on x- axis

5. Shipping and Logistics Analysis.

This analysis focused on understanding the efficiency and cost effectiveness of different shipping modes and impact of order priority on shipping metrics. Different shipping modes exhibit varying average shipping costs and durations, directly impacting overall profitability. For instance, certain expedited shipping options might have higher costs but could lead to higher customer satisfaction due to faster delivery. Reflecting the urgency and associated logistical efforts, the balance between meeting customer expectations for fast delivery and managing shipping costs is crucial.

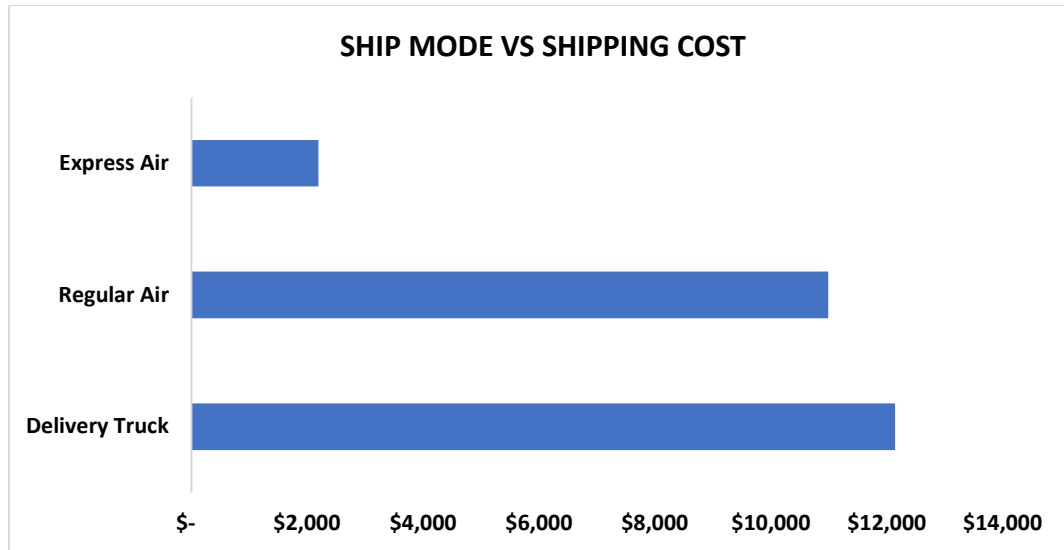


Fig 6. Impact of shipping cost on mode

From the chart above, the analysis shows that Delivery by Truck has the highest shipping cost, follow by Regular Air then Express Air.

6. Regional and State-wise Performance

This analysis provided an in-depth look at sales and profit performance across different regions (East, West, Central and South) for targeted marketing and resource allocations in those regions. Based on this analysis, the East and West regions consistently lead in sales and profit, indicating strong market presence and customer demand, the Central region also shows positive profitability. The south region despite generating sales volume exhibits a significant net loss in profit. This is a critical area requiring immediate investigation into factors such as high operational costs, or unfavorable pricing strategies.

7. Customer Behavior and Segmentation;

Understanding customer groups and their purchasing patterns is essential for developing targeted marketing strategies, personalizing customer experience and optimizing product offerings. All customer segment contributes positively to profit. The Small Business segment, despite lower sales volume demonstrates strong profitability per transaction, suggesting high efficiency in serving this group. Different customer segments show varying preferences for product categories, for example, corporate customers might prefer office supplies and technology, while individual consumers might have broader interests. This insight can inform product stocking and marketing campaigns.

Task 3. Data Visualization and Reporting.

This is the summary of key findings through various data visualizations and an overall reporting overview, enabling quick comprehensive of trends and patterns. The dashboard below summarizes the report.

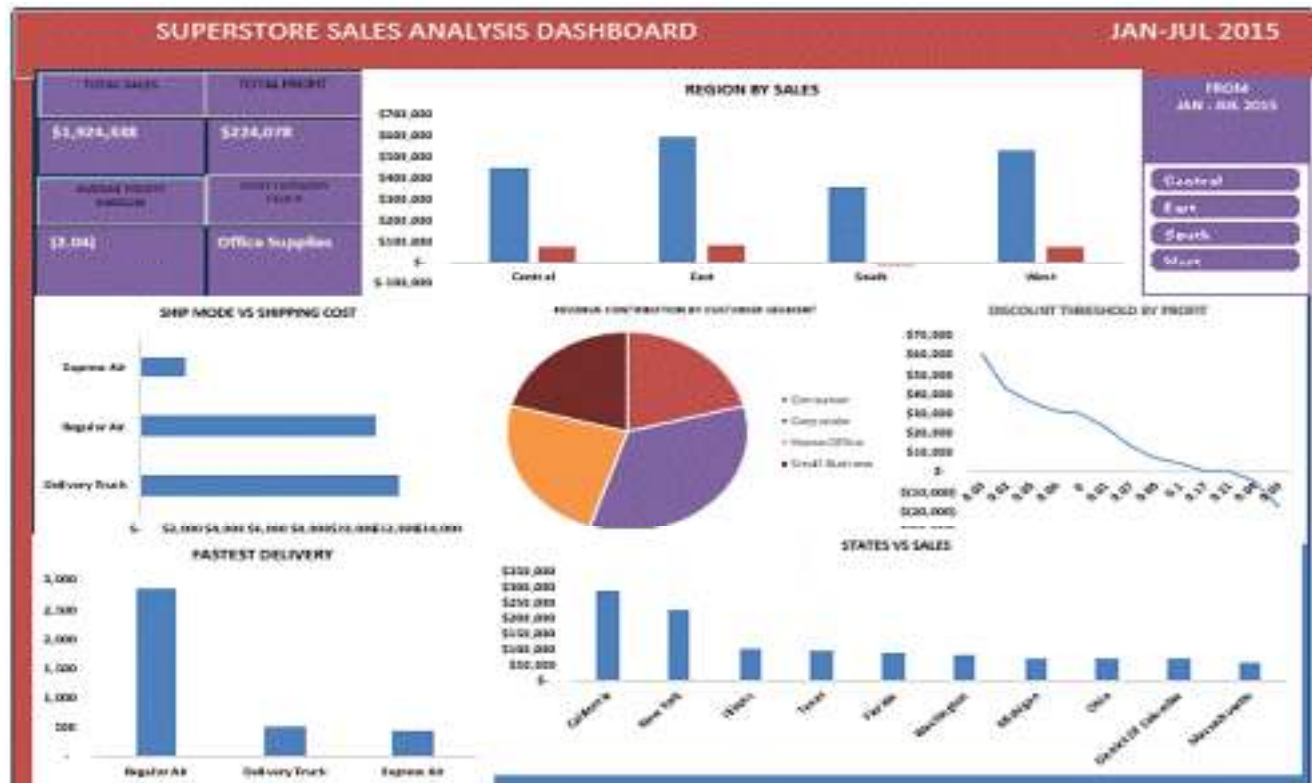


Fig 7. Dashboard

Task 4: Business Insights and Recommendations.

Based on the comprehensive analysis of the SuperStore US 2015 sales data, several key insights have emerged, leading to the following actionable recommendations:

1. Office Supplies: The Hidden Gem; Despite Technology leading in sales, Office Supplies is the most profitable category. This suggests higher profit margins or lower operational costs associated with these products. This is a crucial insight as it highlights that high sales volume does not always equate to the highest profitability.

Increase focus on marketing and promoting Office Supplies, particularly those sub-categories identified as highly profitable (e.g., Binders And Binder Accessories). Explore opportunities to expand the product range within this category.

2. Furniture: High Sales, Low Profit: Furniture, while contributing significantly to sales, yields the lowest profit among the three main categories. This indicates potential issues with its pricing, cost structure, or discount strategies.

Conduct a thorough review of the Furniture category. This should include analyzing procurement costs, pricing strategies, and the impact of discounts. Consider adjusting pricing, negotiating better supplier deals, or reducing discounts on low-margin furniture items.

3. Sub-Category Profitability Discrepancies: There are significant variations in profitability at the sub-category level. While some sub-categories like Binders and Binder Accessories are highly profitable, others such as Rubber Bands, Envelopes, and Pens & Art Supplies are consistently incurring losses. This granular view is essential for precise intervention.

For sub-categories consistently showing losses, conduct a detailed investigation into the root causes. This might involve reevaluating supplier costs, adjusting pricing, improving product quality, or discontinuing products that are not viable.

4. South Region: A Major Concern; The South region is the only one showing a net loss in profit, despite generating sales. This is a critical red flag that requires immediate and in-depth investigation.

Conduct a deep dive into the South region's operations. Analyze local market conditions, competition, operational costs, and pricing strategies. Consider implementing targeted marketing campaigns or adjusting product offerings to better suit the regional demand.

5. Small Business Segment: High Profitability: The Small Business customer segment, while not the largest in terms of sales volume, demonstrates strong profitability. This indicates that transactions with small businesses are highly efficient and yield good returns.

Develop targeted strategies to further engage and expand the Small Business customer segment. This could include specialized product bundles, dedicated customer support, or loyalty programs to capitalize on their high profitability.

6. Strategic Discounting: The analysis on discount impact highlights that excessive discounting can erode profit margins. The identified threshold of 0.17 or 17% for unprofitable discounts is a crucial guideline.

Implement a more strategic approach to discounting. Avoid offering discounts above the 0.17 or 17% threshold unless there is a clear strategic benefit (e.g. clearing old inventory, attracting new customers with high lifetime value). Focus on value-added promotions rather than deep discounts.

8. Shipping Cost Optimization: The relationship between shipping costs and profit needs continuous monitoring.

Regularly review shipping carrier contracts and explore alternative shipping methods to reduce costs without compromising delivery times. Analyze the profitability of orders with high shipping costs to identify opportunities for optimization or to adjust pricing for such orders.

1.6 Limitations for the Project

While this analysis provides valuable insights into the SuperStore US 2015 sales data, it is important to acknowledge certain limitations that may affect the scope and depth of the findings:

1. **Data Scope and Timeframe:** The analysis is based solely on the SuperStore US 2015 dataset. This means the insights are specific to that year and may not fully represent current or future trends. Business dynamics, market conditions, and customer behaviors can change significantly over time, so conclusions drawn from 2015 data might not be directly applicable to subsequent years without further analysis.
2. **Data Granularity:** While the dataset provides detailed information down to the sub-category level, it lacks certain granularities that could offer deeper insights. For example, information on individual product features, manufacturing costs, or specific marketing channels used for each sale could provide a more complete picture of profitability drivers.
3. **Assumptions on Data Cleanliness:** Although the CLEANED DATA sheet was used, the analysis assumes that the data provided is accurate and free from significant errors or biases that might not be immediately apparent. Any underlying data quality issues in the original dataset could potentially impact the reliability of the conclusions.

These limitations suggest avenues for future research and more in-depth analysis, but within the scope of the provided data, this report offers a foundational understanding of the SuperStore's 2015 sales and profit performance.

1.7 Summary

This project has provided a comprehensive analytical exploration of the SuperStore US 2015 dataset, uncovering critical insights into sales performance, customer behavior, and regional trends. Through detailed analysis of shipping and logistics, regional performance, customer segmentation, discount impact, shipping cost influence, and product performance, we have identified key areas of strength and opportunities for improvement.

The findings underscore the importance of a data-driven approach to business strategy. By focusing on optimizing profitable categories like office supplies, re-evaluating underperforming ones like Furniture, addressing regional challenges in the South, and implementing strategic discounting and shipping cost management, the SuperStore can significantly enhance its overall business performance and profitability.

1.8 Conclusion

The SuperStore US 2015 data analysis has yielded valuable insights that can empower the SuperStore to make informed decisions and drive sustainable growth. The recommendations provided are actionable and aim to optimize various facets of the business, from product offerings and pricing strategies to regional focus and customer engagement. Continuous monitoring and adaptation based on these data-driven insights will be key to the SuperStors's ongoing success in the competitive retail landscape.